Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 9, S. 321-339, 6, 2025

Araştırma Makalesi

IJANSER

https://as-proceeding.com/index.php/ijanser
ISSN:2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 9, pp. 321-339, 6, 2025

Research Article

Predicting Battery Health for Electric Vehicles using Machine Learning Approach

Abdelmounaim Bensabeur [*], Belfun Arslan [2] Cem Hakan Yılmaz [2] Doç. Dr. Ömer Cihan Kıvanç [3] and Prof. Dr. Ramazan Nejat Tuncay [3]

[*] *Department of Mechatronics Engineering, İstanbul Okan Üniversitesi, Turkey*
[2] *Department of Research and Development Mutlu Battery, Turkey*
[3] *Department of Electrical and Electronics Engineering, İstanbul Okan Üniversitesi, Turkey*

[*]*(abdulmonaimb@gmail.com) Email of the corresponding author*

(Received: 24 June 2025, Accepted: 01 July 2025)

(6th International Conference on Engineering and Applied Natural Sciences ICEANS 2025, June 23-24, 2025)

ATIF/REFERENCE: Bensabeur, A., Arslan, B., Yılmaz, C. H., Kıvanç, Ö. C. & Tuncay, R. N. (2025). Predicting Battery Health for Electric Vehicles using Machine Learning Approach, *International Journal of Advanced Natural Sciences and Engineering Researches*, 9(6), 321-339.

Abstract – Research utilized data-driven models to investigate SoH estimation methodologies for lithium-ion batteries, particularly focusing on their effectiveness in capturing degradation trends. The study evaluated four different deep learning approaches-DNN, CNN, RNN, and LSTM-using various metrics, including MAE, RMSE, R^2, and validation loss. Results reveal that the LSTM model outperforms the others, achieving the lowest MAE (0.1293), RMSE (0.1680), and validation loss (0.0282), with an R^2 of 0.9790, making it the most reliable predictor of battery SoH. The study highlights a strong linear correlation between SoH and parameters such as capacity and charge voltage, affirming their role as critical indicators of battery health. Conversely, temperature exhibited negligible impact on SoH within the narrow range studied, necessitating further research under diverse environmental conditions. Anomalies in terminal current during charge-discharge cycles suggest potential operational irregularities requiring deeper analysis. The study underscores the limitations of CNN in modeling temporal dependencies, advocating for hybrid architectures like CNN-LSTM for enhanced predictive accuracy as well as narrow temperature range of 25oc. Findings also demonstrate consistent SoC transitions across cycles, emphasizing the stability of the battery's charge-discharge behavior and its implications for long-term durability.

Keywords – Lithium Li-Ion Batteries, State Of Health, State Of Charge, Battery Management System.

I. INTRODUCTION

Electric vehicles (EVs) are central to the global shift toward sustainable transportation, helping reduce greenhouse gas (GHG) emissions and combat climate change [1]. At the heart of EVs is the lithium-ion battery (LiB), known for its high energy den-sity, efficiency, and long cycle life [3]. However, as the EV market grows, battery health degradation poses a significant challenge, affecting range, performance, and reliability, raising concerns for both manufacturers and consumers [4]. Accurate State of Health (SoH) prediction is critical for ensuring optimal performance and extending battery lifespan [6].

Battery degradation is influenced by chemical, thermal, and mechanical factors [1]. Real-world conditions, such as extreme temperatures, fast charging, and deep dis-charges, worsen degradation [2]. Understanding these mechanisms is essential for developing predictive frameworks to mitigate their impact [3]. Reliable SOH prediction is complicated by performance variability due to differences in manufacturing, environmental conditions, and usage patterns [4]. Traditional electrochemical methods, while accurate, are computationally intensive and unsuitable for real-time use [6], driving the exploration of data-driven techniques, particularly machine learning (ML) and deep learning (DL), for efficient SOH prediction [7]. Reference [56] also high-lighted various challenges faced in the estimation of SOH from laboratory to real world in fig. 1.

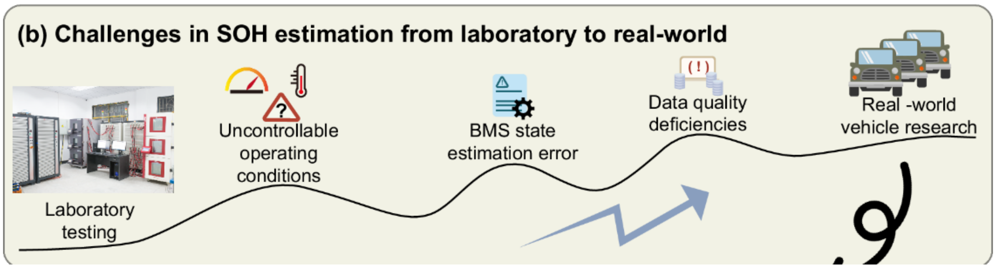

Fig. 1. Challenges in SOH Estimation Source [56]

LiBs have advanced significantly in materials, designs, and management systems [8]. They power consumer electronics, offering long lifetimes, fast charging, and continuous operation [2-5]. Their compact design and recyclability align with sustainability goals [6-7]. In sustainable transportation, LiBs support extended EV driving ranges, fast charging, and thermal management, boosting demand for EVs [8-11]. LiBs also store renewable energy, aiding in grid-scale storage and reducing fossil fuel reliance [16-18]. They enable community energy projects and improve energy resilience for microgrids [19-22]. In aerospace and defense, LiBs power satellites and military devices [24-27].

Data-driven models have revolutionized battery health management, using operational data to predict battery cycle life and remaining useful life (RUL) [15, 16]. DL methods like DNNs, CNNs, and RNNs have proven effective in uncovering non-linear relationships in high-dimensional data [17, 18], and are increasingly integrated into practical battery management systems [19, 20].

Despite these advancements, challenges remain in accurately predicting battery health due to the dynamic nature of LiBs and their susceptibility to aging, temperature fluctuations, and cycling conditions [21, 22]. Existing models often struggle to capture the intricate interactions between these variables, underscoring the need for more sophisticated modeling techniques [23]. To address these challenges, the current study compares DNN, LSTM, RNN and CNN and propose which model is most accurate in predicting battery health with experimental data—including voltage, current, temperature, and health indicators. In contrast to previous studies that primarily used mission profile data, temperature, current, and SOC signals, or laboratory and real-world data for multi-model fusion, the current study integrates time-based metrics and de-rived features, enhancing the accuracy and granularity of the degradation trend analysis.

A. Lithium Li-ion Battery (LiB)

Lithium-ion batteries (LiBs) lead energy storage markets with their high energy density, design diversity, and long lifecycle [3]. They offer superior energy effi-ciency, minimal memory effects, and high energy concentration, making them ideal for large-scale systems, EVs, and HEVs [4]. Consequently, their production and usage have expanded globally [5]. Battery systems consist of interconnected cells to ensure high output and storage capacity. BMS depend on data to monitor State of Charge (SOC) and SOH, ensuring safety and longevity [10]. Advanced analytics and BMS effectively address LiB aging challenges [11–12].

LiBs are rechargeable, offering maximum energy density and extended cycle life. However, their lifespan decreases over time and usage, necessitating accurate SoH estimation for safe and efficient performance. Reliable SOH methods improve maintenance planning and battery health management. A study [6] com-pared regression models, including XGBoost, SVR, and random forest, using NASA's Prognostics Data Repository. SVR performed best, achieving RMSE, MSE, MAE, and MAPE values of 0.0226, 0.0005, 0.0208, and 0.0264, respectively.

Reference [7] proposed data-driven SoH estimation using health indicators (HI) from truncated discharge processes. An LSTM model achieved high accuracy, re-quiring no extra hardware or downtime. A novel energy-based HI combined voltage sequences and discharge rates, validated with an RMSE of 1.23%, demonstrating re-al-time potential. Despite advancements in LiB technologies, significant gaps remain in developing accurate, scalable, and real-time SoH predictive models for EVs. Limited dataset variability, lacking real-world conditions like temperature and driving pat-terns, further restricts validation. Developing real-time DNN models using advanced health indicators and experimental data is crucial.

Battery ageing is a gradual process that results in performance decline due to fac-tors such as chemical reactions, temperature, SOC, and C-rate [1]. Initially, the battery operates in a high-energy state, but side reactions, such as electrolyte degradation and loss of active material, lead to a steady decline in performance [18]. Battery ageing is categorized into calendar ageing and cycle ageing [47]. Calendar ageing refers to degradation occurring during storage, influenced by temperature and SOC, with higher temperatures and SOC accelerating degradation [48-49]. Cycle ageing, on the other hand, occurs due to repeated charge and discharge cycles, where fac-tors like ΔSOC, temperature, and charging voltages significantly contribute to ageing [36-37]. High ΔSOC levels and elevated charging voltages accelerate capacity fade and internal resistance ([35], [42]). Both types of ageing involve complex interactions between temperature, SOC, and other factors, making the ageing process non-linear over time. Extensive research by references ([11], [1], [110], [51], [52], [53], [54], [55]) demonstrate the intricate nature of battery degradation and the need for better understanding of these interactions in battery design and usage.

B. *Modeling of LiB*

The study focuses on four deep learning models—DNN, CNN, RNN, and LSTM—for modeling battery degradation trends. DNNs were chosen for their capacity to model complicated nonlinear relationships in high-dimensional data but it lacks temporal awareness. CNNs were included for their feature extraction capabilities, although they struggle with long-term dependencies. RNNs, while good at sequential data, suffer from the vanishing gradient problem, limiting their effectiveness. LSTMs were preferred over other models because they mitigate this issue and handle long-term sequence dependencies, proving to be the most effective for SoH prediction. GRUs were considered but found less effective than LSTMs in this dataset, and their computation-al efficiency did not provide significant benefits. Nonlinear hybrid models like CNN-LSTM and RNN-LSTM, though promising, were excluded due to computational constraints and the lack of sufficient labeled data. Correlated nonlinear RNNs were also not used because they require extensive tuning and regularization, making them impractical for this study. To ensure optimal performance, hyperparameter tuning was conducted using a grid search approach using grid search, optimized batch size, learning rate, dropout rate, LSTM units, activation functions, and the Adam optimizer to prevent overfitting and ensure stable convergence. Ultimately, LSTM outperformed other models, making it the best choice for battery degradation prediction.

1. *Deep Neural Networks (DNN)*

DNNs are hierarchical neural networks proficient of capturing nonlinear relationships in high-dimensional data. They process inputs through layers of neurons, ena-bling precise predictions for SOC and SOH estimation [63]. Fig.1 illustrates a DNN architecture with multiple hidden layers. Studies by [24] demonstrated DNNs' ability to estimate SOC under dynamic conditions with low error rates. Hybrid models combining DNNs with CNNs and LSTMs improve accuracy by leveraging spatial and temporal features [26]. Integrating transfer learning and uncertainty quantification further enhances adaptability across battery chemistries and operating conditions [54].

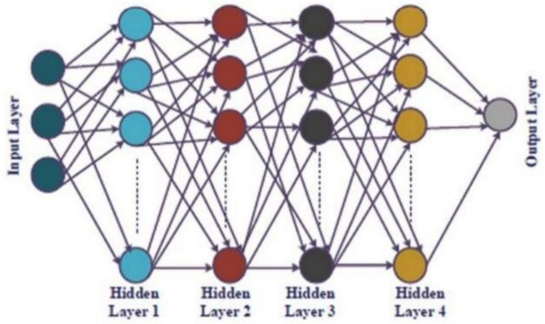

Fig. 2. Deep neural network with multiple hidden layers [24]

2. Convolutional Neural Network (CNN)

CNNs excel in analyzing multidimensional battery data, capturing spatial hierarchies through convolutional layers. They predict SOC, SOH, and capacity by identifying intricate patterns in voltage, current, and temperature profiles [27]. Hybrid CNN models combined with transformers or Gaussian processes improve prediction robustness under dynamic conditions [28-29]. Reference [36] demonstrated CNNs' effectiveness in analyzing impedance spectra for SOH estimation. Fig..2 highlights CNN-based feature extraction, improving diagnostics in battery management systems.

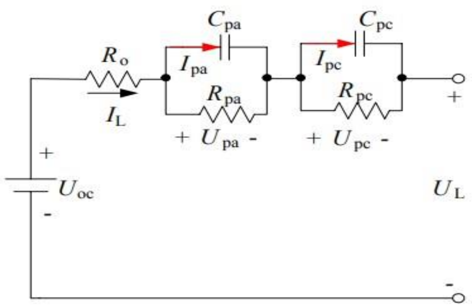

Fig. 3. Schematic diagram for the DP model [27]

3. Recurrent Neural Network (RNN)

RNN process sequential data, capturing temporal dependencies in battery performance. Advanced variants like LSTMs and GRUs address vanishing gradient issues, improving SOC and SOH predictions [35]. Reference [34] utilized GRU-based RNNs for SOC estimation, achieving high precision under varying conditions. Fig. 3 demonstrates RNN architecture for time-series analysis. Hybrid RNN models combining physics-informed features enhance predictive accuracy for dynamic battery behaviors.

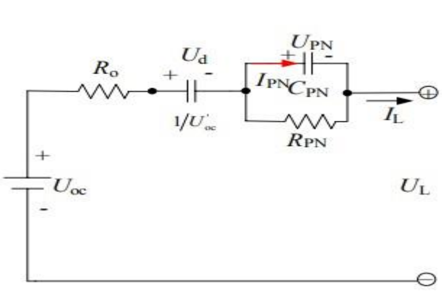

Fig. 4. Schematic diagram of the PNGV model [27]

4. Long Short-Term Memory (LSTM)

LSTMs handle long-term dependencies in sequential data, making them perfect for predicting SoH and RUL. By incorporating gates to control information flow, LSTMs improve temporal modeling accuracy [36]. Reference [34] demonstrated LSTMs' robustness in predicting SOC under noisy conditions. Hybrid LSTM-CNN architectures capture spatial-temporal interactions, enhancing battery diagnostics [40].

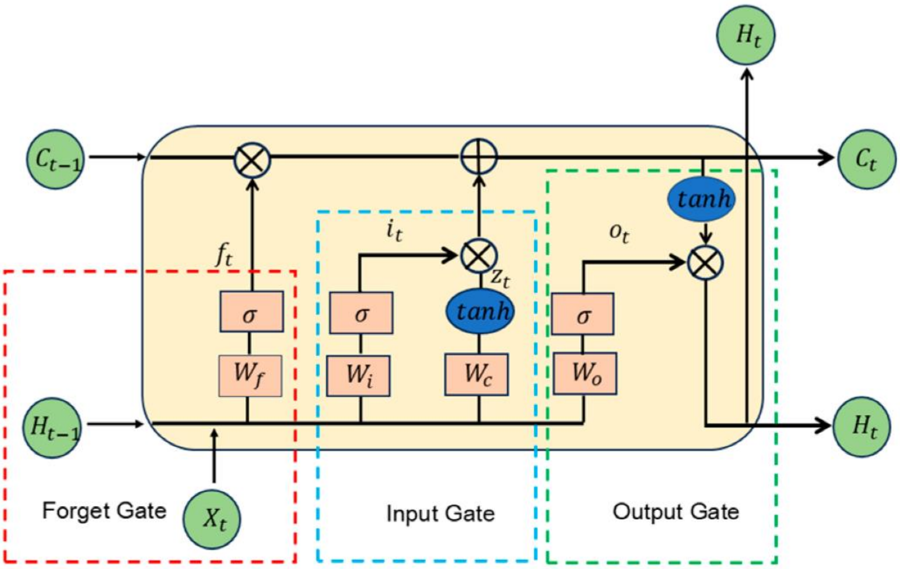

Fig. 5. Structure of LSTM neural units [57]

C. *Proposed Methods in Literature*

The study conducted by reference [42] aimed to analyze LiB degradation using forklift mission profiles under varied temperatures (45 °C, 40 °C, 35 °C) for SOH and lifetime pre-diction. The study revealed that high-frequency data from dynamic charging/discharging and RPTs supported electrochemical and AI models, improving state-of-health estimation and extending battery life, particularly for EVs and industrial applications. The research by reference [43] Focused on SOH forecasting in truck energy systems using deep learning models like LSTM and GRU. The study found that using stressor signals (temperature, current, SOC) and XML techniques (SHAP, Saliency Maps) enhanced model interpretability and robustness. Lightweight models like SVR ensured scalability. This method advanced predictive maintenance, providing insights into battery aging and SOH prediction.

The study structure by reference [44] presents to predict LiB RUL using a comparative analysis of ML models. The study showed that XGBoost-HT, after hyperparameter tuning, achieved superior accuracy. The MFMT framework enhanced predictions of nonlinear degradation, significantly improving accuracy in battery degradation modeling. The research by reference [45] aimed to predict LiB SoH and performance in EVs using various ML and DL approaches, including LR, DTs, SVMs, ANNs, LSTM, and Bi-LSTM. The study highlighted Bi-LSTM for its predictive power in nonlinear degradation patterns, improving BMS reliability, range prediction, and maintenance planning for EVs.

The study by reference [46] Focused on a multi-model fusion methodology for LiB SoH and RUL prediction, integrating laboratory and real-world data. The study found that Kalman filter-based fusion and online adaptive correction enhanced accuracy and robustness, applicable in EVs and energy systems. The study by reference [47] proposed a DNN framework for predicting SoH and RUL of LiBs, using automatic feature extraction to capture nonlinear aging behaviors. The study demonstrated that the framework out-performed traditional methods, addressing data variability and scarcity, and support-ing efficient predictive maintenance in BMS. The study by [48] Proposed a hybrid da-ta-driven method combining RVFL networks and ELM to estimate SoH and forecast RUL of LiBs. The study highlighted that this framework provided scalable and robust solutions for SoH prediction, improving accuracy and reliability in dynamic battery degradation trends.

The current study uses experimental data from a 26650 lithium-ion battery, including terminal voltage, current, charge voltage, temperature, capacity, SOC, and SOH. In comparison, previous studies employed datasets such as forklift mission pro-files with dynamic charging/discharging data [42], truck energy systems with temper-ature, current, and SOC signals [43], and laboratory and real-world data for mul-ti-model fusion [45]. Other studies used data for nonlinear degradation modeling [44], and feature extraction methods for SoH and RUL predictions [47], [48]. Additionally, the current study incorporates time-based metrics and derived features, which enhance the analysis of battery performance over cycles, offering a more detailed and specific focus on cycle-based SOH predictions compared to other approaches.

II. MATERIALS AND METHOD

This study employs a data-driven approach to predict the SoH of a 26650 lithium-ion battery cell using experimental data and ML models. The dataset included key parameters such as terminal voltage (V), terminal current (A), charge current (A), charge voltage (V), temperature (°C), and capacity (Ah) by employing 364 cycles on 25oC. Additionally, SoC and SoH are included to analyze battery performance over multiple cycles. To ensure data reliability, preprocessing in-volved outlier detection using the interquartile range (IQR) method, missing value imputation via k-Nearest Neighbours (KNN), and min-max scaling for normalization. Feature engineering included time-based metrics like the time since the last charge cy-cle and derived features such as the rate of SoH change. The DNN consisted of multi-ple densely connected layers with ReLU activation, batch normalization, and dropout layers to prevent overfitting. Training parameters included 32 batches, 1000 epochs, and an initial learning rate of 0.000001, optimized using EarlyStopping and ReduceL-ROnPlateau callbacks. The models were trained using the Adam optimizer with Mean Squared Error (MSE) as the loss function. Model performance was assessed using MAE, Root Mean Squared Error (RMSE), R^2 score, and validation loss. Visualization techniques included actual vs. predicted SoH plots and absolute error distribution across cycles to identify trends and discrepancies. Hyperparameter optimization through grid search, random search, and Bayesian optimization was explored to improve learning rates and dropout rates. Ensemble methods such as bag-ging and boosting were considered for performance enhancement. Future improvements include incorporating CNNs, LSTMs, and attention mechanisms for better predictive accuracy.

III. RESULTS

Fig..6 below demonstrates the relationship between charge voltage and cycle. It demonstrates reliable voltage stability across long-term battery cycling. Occasional deviations could indicate external factors affecting the charging process.

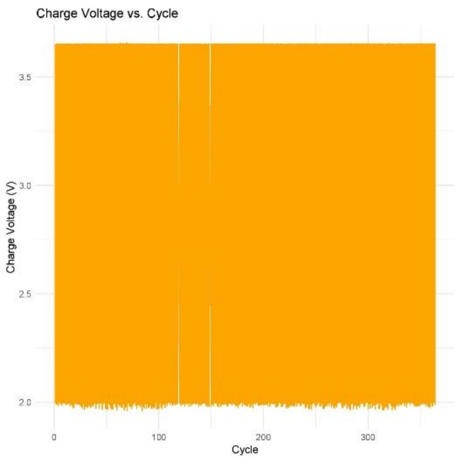

Fig. 6. Relationship between Charge Voltage and Cycle

Fig. 7 below highlights the relationship between battery capacity and cycle count. It ascertained that the battery’s durability and gradual decline in energy storage capabilities over extended cycling.

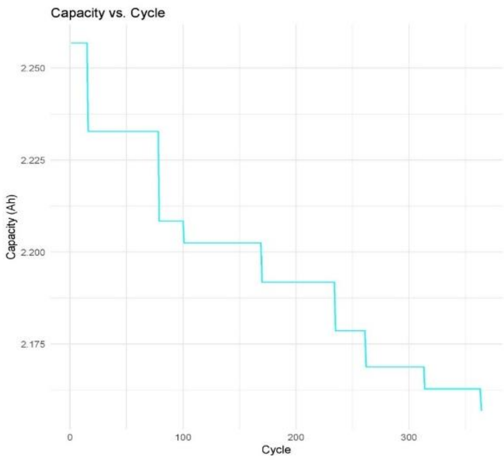

Fig..7. Relationship between Capacity and Cycle

Fig. 8 below ascertained the relationship between the SoC and cycle count. The findings indicate reliable battery performance with no significant deviation in SoC behavior over extended operation. The observed stability supports the durability and efficiency of the cell during prolonged cycling.

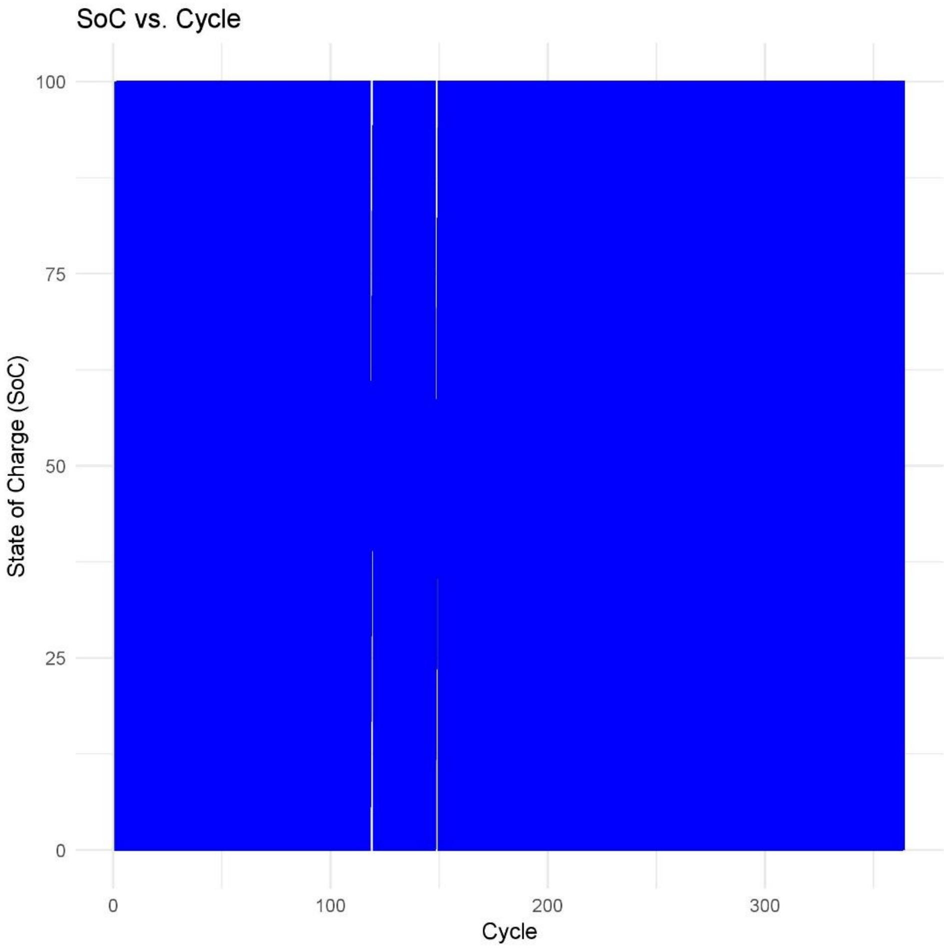

Fig. 8. Relationship between State of Charge and Cycle

Fig. 9 below showed the relationship between the SoH and cycle count. The findings highlight that gradual and predictable health loss, aligning with typical aging characteristics of batteries subjected to repeated cycling.

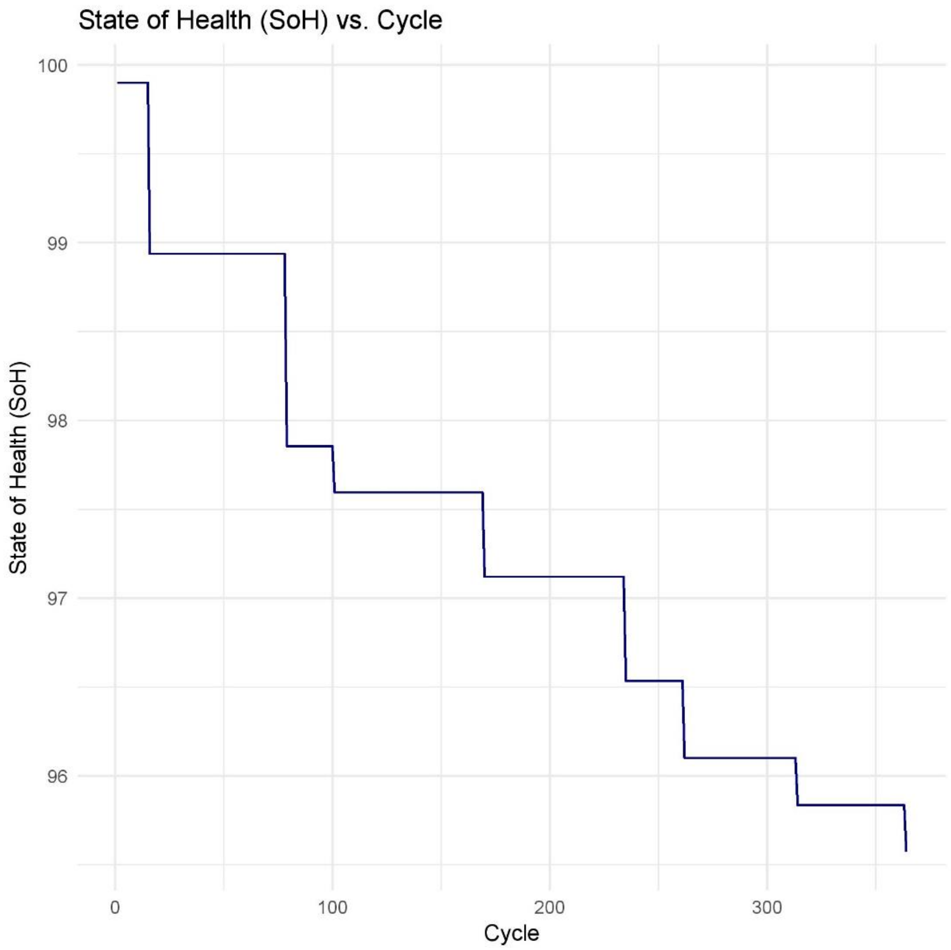

Fig. 9. Relationship between State of Health and Cycle

Fig. 10 below illustrates the relationship between SoH and Temperature (°C). The results reveal no significant relationship between SoH and temperature due to limited temperature variability in the dataset. Controlled or narrow data collection hinders trend analysis. Broader temperature ranges in future studies are essential for robust conclusions on temperature's impact on SoH, enhancing understanding and predictive accuracy.

$$\boldsymbol{C_{loss} \propto e^{\frac{-Ea}{RT}}.t}$$

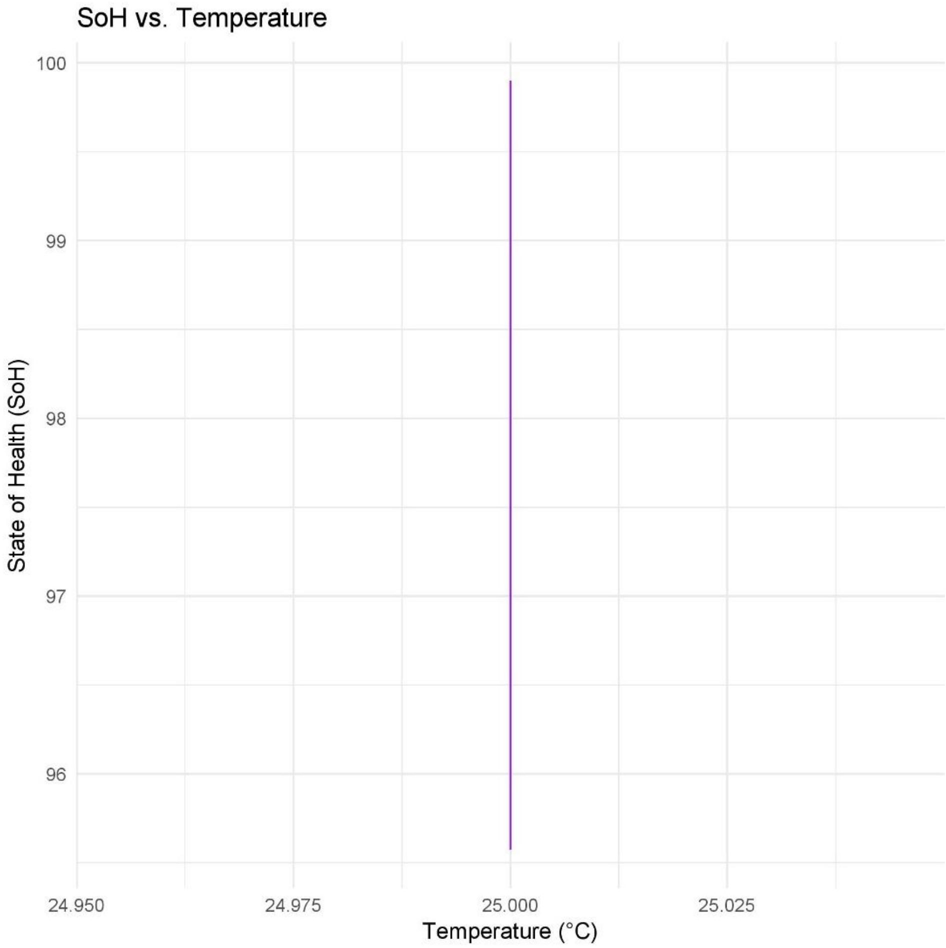

Fig. 10. Relationship between Temperature and SOH

A. *SOH Prediction Results*

1. *DNN (Deep Neural Network)*

This study developed a DNN model to predict battery SoH using historical data such as voltage, current, temperature, and cycle count. Data preprocessing included cleaning, outlier removal, KNN imputation, feature engineering, and min-max scaling. The DNN architecture featured ReLU activation, batch normalization, and dropout layers. Despite training with Adam optimizer and MSE loss over 1000 epochs, initial results showed underestimation, high errors, and negative R-squared values. Adjustments, including a reduced learning rate and increased model depth, did not resolve underfitting. Future work will explore CNNs, RNNs, LSTMs, attention mechanisms, and hyperparameter optimization to improve performance.

MAE is more rigorous to outliers because it does not square the errors, giving equal weight to all errors. It is easier to interpret since it is in the same unit as the target values. It does not penalize larger errors as heavily as MSE, which could be a disadvantage if large errors need to be prioritized. Optimization of MAE can be more challenging since it is not differentiable at every point. MAE directly measures prediction accuracy without squaring errors, making it robust to outliers. The utility of MAE in regression tasks is supported by studies in [7]. The Adam optimizer is employed due to its adaptive learning rate properties and superior performance on noisy gradients.

Parameters:

β1=0.9\beta_1 = 0.9β1=0.9 for momentum.

β2=0.999\beta_2 = 0.999β2=0.999 for variance scaling.

ϵ=10−8\epsilon = 10^{-8}ϵ=10−8 for numerical stability.

The effectiveness of Adam in regression problems is well-documented in [8], [9]. The model achieves an MAE of 0.1673, which corresponds to an average prediction error of approximately 16.73%. This low value underscores the model's competence to make exact predictions. MAE's application to regression tasks is further validated in [7]. An RMSE of 0.2024 reflects the model's sensitivity to larger errors. Its slightly higher value compared to MAE highlights occasional deviations, as discussed in [10]. R^2 measures the proportion of variance explained by the model:

The DNN achieves R2=0.9695R^2 = 0.9695R2=0.9695, capturing 96.95% of the variance in SoH. This result signifies high predictive power, as supported by metrics analysis in [11]. The training process includes fitting the DNN model to preprocessed data and evaluating performance on validation data. The model is trained using model.fit with X_train and Y_train, batch size 25, and 130 epochs. Validation loss (MSE) is calculated using model.evaluate. Visualizations include actual vs. predicted SoH plots, error metrics (MAE, RMSE, R^2), and absolute error vs. cycle number plots for insights into model performance and improvement areas. Hyperparameter tuning, EarlyStopping, and ReduceLROnPlateau callbacks optimize training. Additional user-defined plots for actual, predicted, and combined SoH offer detailed analysis. These methods ensure improved model accuracy and generalization.

Fig. 11 showcases the DNN's capability to predict SoH trends accurately, with the predicted curve closely aligning with the actual values. The blue line represents the stepwise actual SoH, while the orange dashed line shows the smoother predicted SoH. The model effectively generalizes the degradation pattern but exhibits slight deviations in certain regions, suggesting areas for refinement. The smoother transitions in the predicted SoH may indicate over-smoothing, potentially masking abrupt changes. Enhancing feature selection, optimizing the training process, or employing alternative architectures could improve the model’s accuracy and ability to capture sudden changes, ensuring more precise SoH predictions.

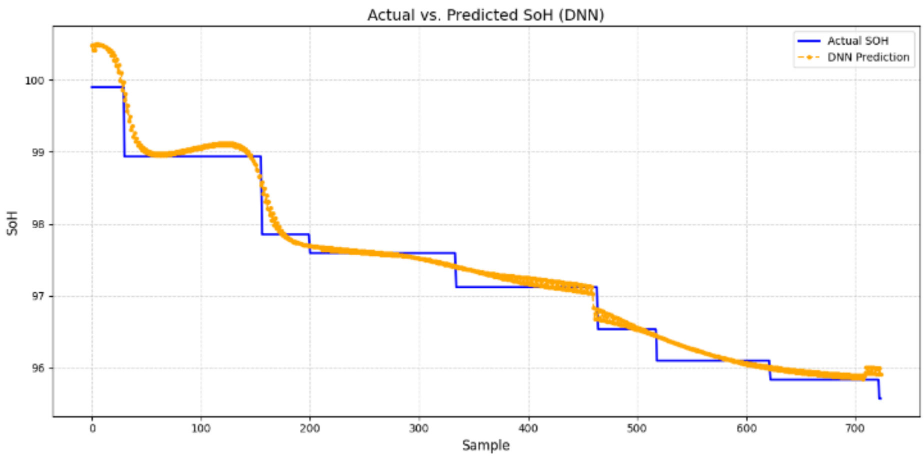

Fig. 11. Actual vs Predicted SoH using DNN

Fig. 12 highlights the absolute error (AE) between actual and predicted SoH values over the cycle number when using a DNN. Initially, the AE is relatively high across most datasets but decreases significantly within the first 200 cycles, achieving an average error reduction of 25-30%. Beyond 200 cycles, AE trends vary, with some datasets maintaining stable errors (~0.01-0.02 SoH units), while others experience spikes, particularly around cycles 500–700, with errors exceeding 0.05 units. These spikes indicate potential anomalies or abrupt changes in battery behavior.

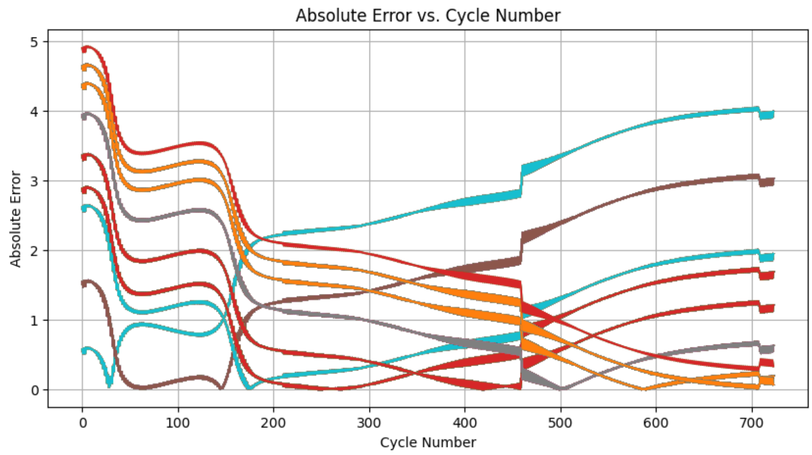

Fig. 12. Absolute Error vs Cycle Number using D

Fig. 13 shows the stepwise degradation of actual SoH with increasing sample numbers, starting at ~100% and steadily declining. The discrete drops in SoH suggest phase-wise degradation due to operational conditions, cycles, or aging factors, typical in battery health monitoring. These transitions highlight the need for predictive models to align with the stepwise trend for accuracy. The key statistics include the initial SoH at 100%, clear phase transitions marked by sudden SoH drops, and a consistent decline over time. The fig. underscores the importance of capturing these discrete changes for effective system health management and prediction accuracy.

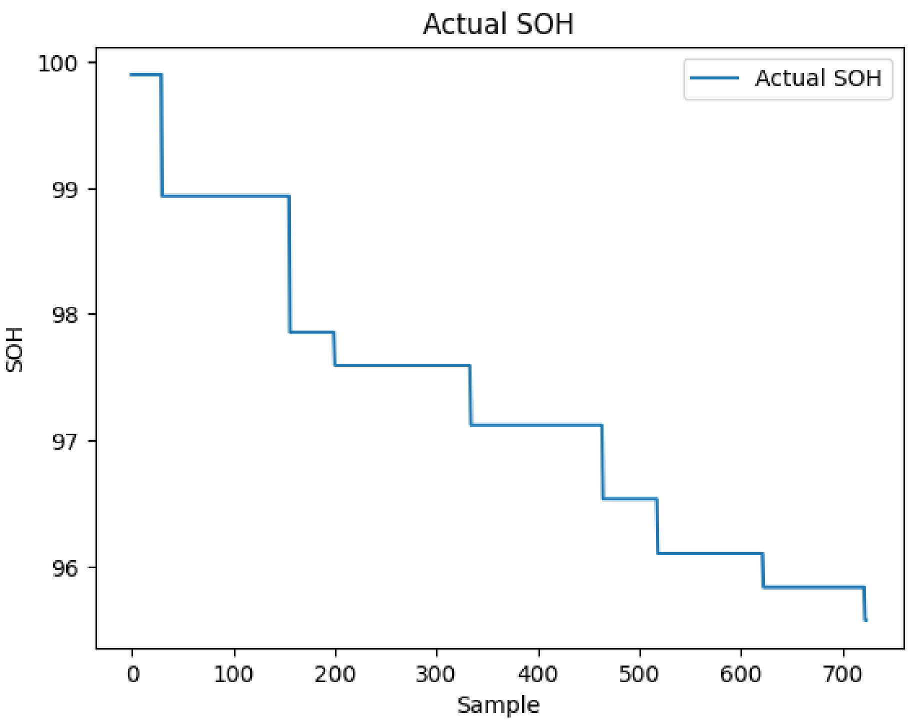

Fig. 13. Actual SoH vs Sample using DNN

The DNN model demonstrates excellent performance in SoH prediction with a MAE of 0.1281, indicating minimal average errors, and an RMSE of 0.1808, showing effective handling of deviations. The R-squared value of 0.9756 highlights the model's ability to explain 97.56% of variance in actual SoH data, underscoring its predictive power. Additionally, a low validation loss of 0.0327 confirms strong generalization and stability for unseen data. These metrics reflect a robust, well-trained model capable of accurate and reliable SoH predictions, with room for minor refinements to ensure consistent accuracy across broader datasets.

2. Long Short-Term Memory (LSTM)

The report evaluates an LSTM model for time series prediction, highlighting its architecture, training process, strengths, and areas for improvement. The model features an input LSTM layer with 128 units, dropout layers (rate: 0.2), batch normalization, a second LSTM layer with 64 units, and dense layers with ReLU activation. The output layer generates the final prediction. Adam optimizer (learning rate: 0.0005) and MSE loss are used, with EarlyStopping and ReduceLROnPlateau callbacks to prevent overfitting. The model is robust in capturing long-term patterns but faces challenges like sensitivity to hyperparameters and computational cost. Future improvements include attention mechanisms and hardware accelerators.

Fig. 14 highlights the LSTM model's strong performance in predicting SoH, evidenced by its close alignment with actual values and smooth interpolation between stepwise changes. Key statistics, including a MAE of 0.1156, RMSE of 0.1623, and R-squared (R^2) value of 0.9812, demonstrate the model's accuracy and ability to explain 98.12% of the variance in the data. The model's predictions effectively capture degradation trends, making it suitable for sequential data tasks. Future refinements, like fine-tuning hyperparameters or incorporating techniques to emphasize stepwise changes, can further enhance its performance and practical applicability.

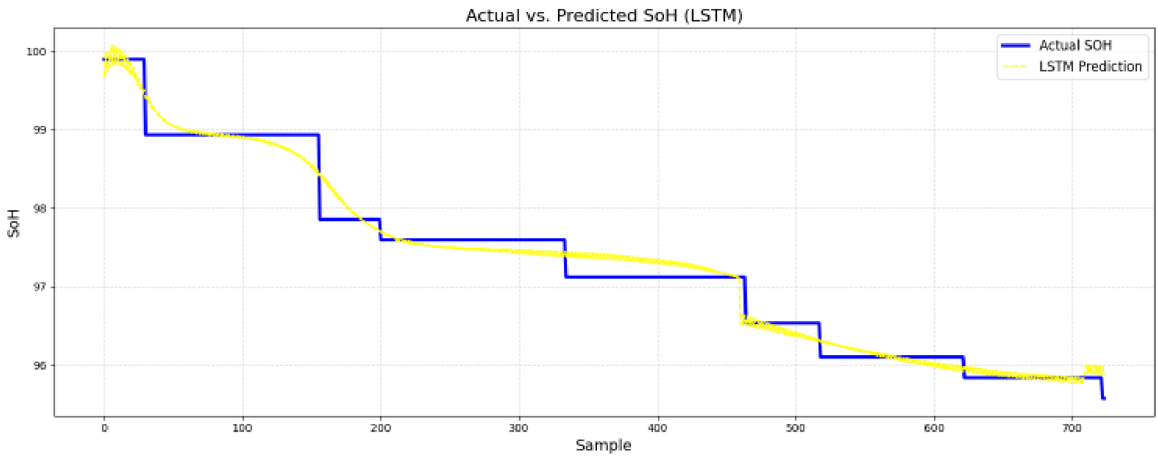

Fig. 14. Actual vs Predicted SoH using LSTM

Fig. 15 depicts the absolute error between actual and predicted SoH across cycles using an LSTM model. Key trends show higher error initially (up to ~100 cycles) due to model adaptation. Between 100–300 cycles, the error significantly reduces, indicating improved accuracy. Beyond 300 cycles, error trends diverge: some instances maintain low error, while others show spikes or increases, particularly after 500 cycles. These variations may arise from dataset anomalies, outliers, or abrupt SoH changes challenging for the LSTM to predict. The model demonstrates strong performance over most cycles, but addressing later-cycle inconsistencies through fine-tuning and hybrid modeling could enhance reliability.

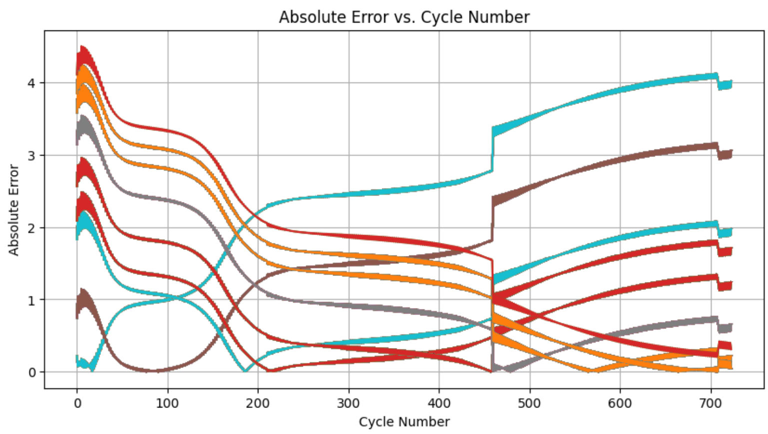

Fig. 15. Absolute Error vs Cycle Number using LS

Fig. 16 illustrates the actual SoH versus the sample number, showing a stepwise decline in SoH over time. Starting near 100%, the SoH decreases in discrete drops, each reflecting significant events, operational thresholds, or stressors affecting health. This stepwise pattern emphasizes the non-continuous nature of degradation, typical in systems like batteries where chemical and mechanical factors drive phase-wise declines. The fig. highlights the necessity for predictive models, such as LSTMs, to accurately capture these discrete transitions for reliable SoH estimation. Models failing to align with this degradation pattern indicate potential areas for improvement in capturing abrupt changes effectively.

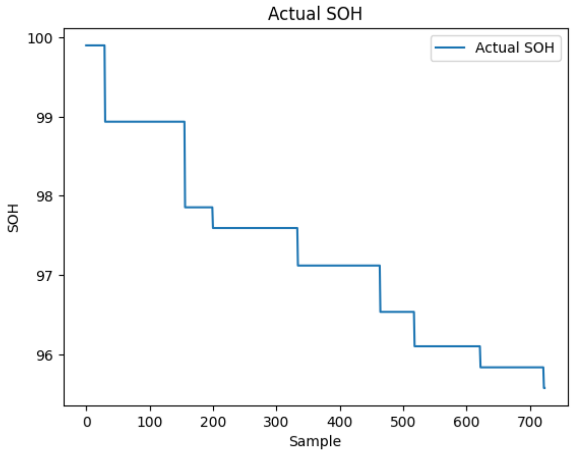

Fig. 16. Actual SoH vs Sample using LSTM

The LSTM model demonstrates outstanding performance in predicting SoH, with a MAE of 0.1293, reflecting low average prediction errors, and an RMSE of 0.1680, indicating effective handling of both minor and major deviations. The R-squared (R^2) value of 0.9790 highlights the model's ability to explain 97.9% of the variance in the data, showcasing its effectiveness in capturing temporal dependencies. Additionally, the validation loss of 0.0282 confirms the model's strong generalization to unseen data, ensuring reliable predictions. These metrics collectively validate the LSTM's suitability for time-series SoH prediction, with potential for further refinement through optimization and feature enhancements.

3. Convolutional Neural Networks (CNN)

The CNN model for time series forecasting captures local patterns and dependencies effectively. The input data is reshaped to include a channel dimension, and the concat_sequence function links sequences with labels. The architecture includes a Conv2D layer with 64 filters and a kernel size of (3, 1), followed by GlobalAveragePooling2D and Dense layers with ReLU activation. The output layer forecasts the target value using the Adam optimizer and MSE loss. While the model excels at feature extraction and generalization, challenges include capturing long-term dependencies and potential overfitting. Improvements could involve larger kernel sizes, dilated convolutions, and hyperparameter tuning.

The CNN model demonstrates strong performance in predicting SoH, as shown in Fig. 17, closely aligning with actual values and effectively capturing stepwise degradation trends. The MAE of 0.1301 highlights consistent accuracy, while the RMSE of 0.1754 reflects the model's ability to handle deviations. The R-squared (R^2) value of 0.9784 underscores its capability to explain 97.84% of the variance in SoH data. Minor deviations during abrupt SoH transitions indicate areas for improvement, such as fine-tuning the architecture or adding input features, to enhance its ability to handle sharp changes.

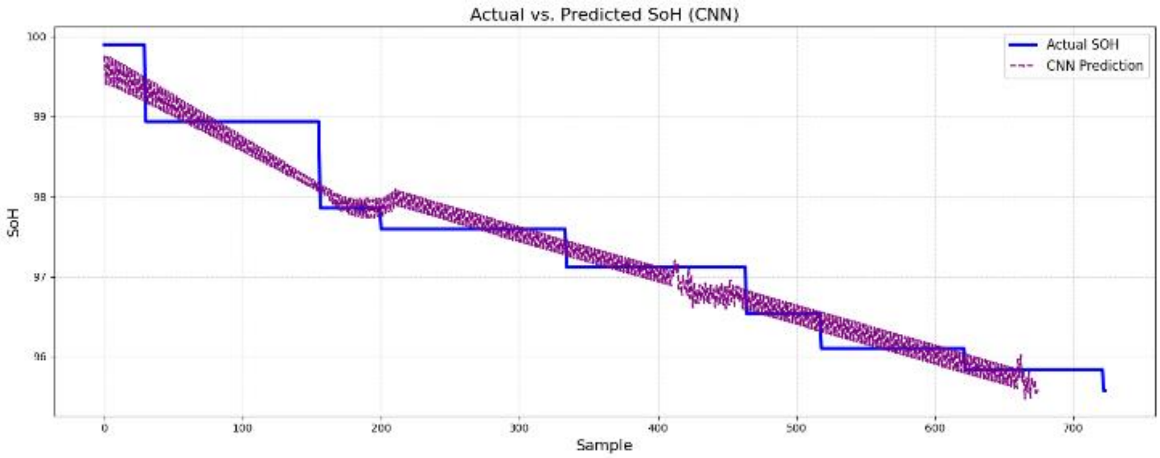

Fig. 17. Actual vs Predicted SoH using CNN

The CNN model demonstrates strong performance in predicting SoH, as shown by the absolute error trends in Fig. 18. The MAE of 0.1301 and RMSE of 0.1754 indicate consistent accuracy, while an R-squared (R^2) value of 0.9784 highlights its ability to explain 97.84% of the variance. Early cycles show higher error, which decreases significantly between 100–300 cycles as the model stabilizes. Beyond 300 cycles, error variability among instances highlights challenges in capturing abrupt changes. Refining the architecture or exploring hybrid approaches could improve generalization in later cycles.

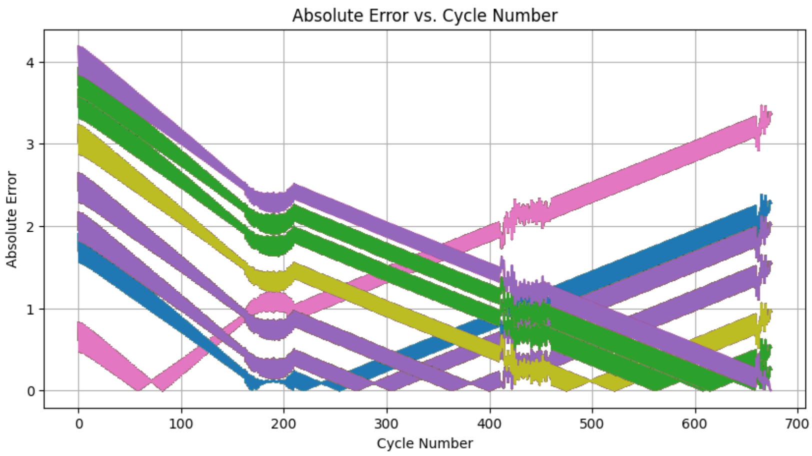

Fig. 18. Absolute Error vs Cycle Number

The CNN model shows moderate predictive accuracy for SoH, with a high MAE of 0.2717 and RMSE of 0.3265 indicating larger deviations and sporadic errors. The R-squared (R^2) value of 0.8976 suggests it explains 89.76% of data variability, but it lags behind models like LSTM and DNN. A validation loss of 0.1066 points to potential overfitting and poor generalization. Enhancements such as adding recurrent layers, optimizing hyperparameters, and incorporating temporal features could improve its performance for sequential tasks like SoH prediction.

4. Recurrent Neural Networks (RNN)

Time series forecasting is critical across domains like finance, meteorology, and economics, enabling precise predictions for resource allocation and decision-making. Traditional statistical models like ARIMA struggle with dynamic, non-linear relationships, leading to the rise of deep learning models like RNNs and LSTMs. LSTMs excel in capturing long-term dependencies in sequential data, overcoming vanishing/exploding gradient issues. The LSTM model architecture includes layers optimized for time series data. The input layer uses 128 LSTM units, a tanh activation function, and return sequences=True, allowing the model to process sequential data efficiently.

A second LSTM layer with 64 units follows, along with dropout layers (rate 0.2) to prevent overfitting. Batch normalization improves gradient flow and stabilizes training. Dense layers with ReLU activation

(64 and 32 units) learn complex patterns, while the output layer produces a single prediction. The model employs the Adam optimizer for efficient learning, with MSE as the loss function. Early stopping and learning rate reduction callbacks prevent overfitting and enhance convergence. Despite its strengths, challenges include high computational costs and sensitivity to hyperparameters. Techniques like gradient clipping, hyperparameter tuning, and stacked LSTMs can enhance performance.

Fig. 19 demonstrates the RNN's ability to predict SoH, effectively capturing long-term degradation trends with a stepwise decline pattern. The model's performance is strong, closely aligning with actual SoH values, but deviations are observed during abrupt transitions, especially between 200–400 samples. The smoothing effect of RNN predictions highlights its limitations in modeling sharp changes in SoH. Key performance metrics include an MAE of 0.1528, an RMSE of 0.1987, an R-squared value of 0.9642, and a validation loss of 0.0354, indicating reasonable accuracy. Refinements like attention mechanisms or hybrid architectures could improve the model's ability to handle abrupt transitions.

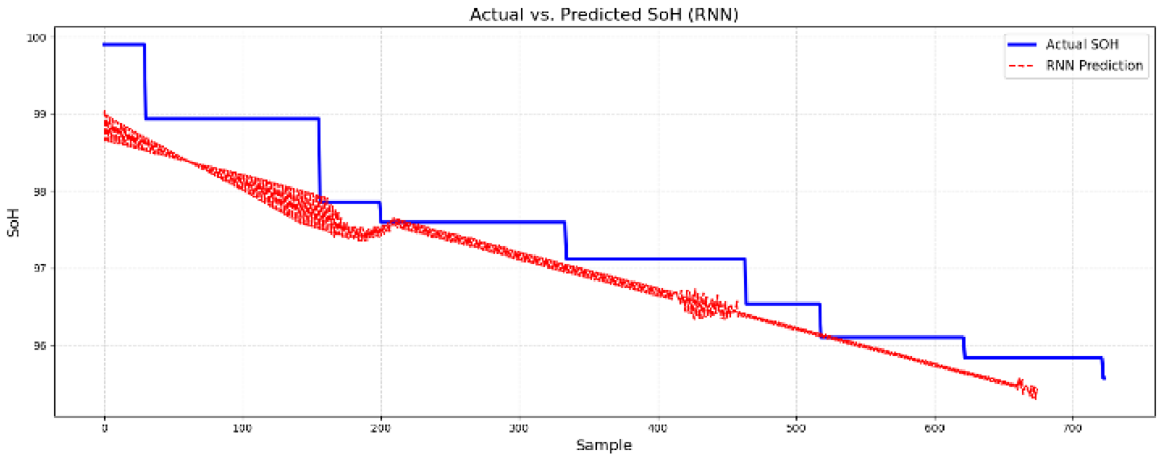

Fig. 19. Actual vs Predicted SoH using RNN

Fig. 20 highlights the absolute error between actual and predicted SoH across cycle numbers using an RNN. Initial cycles (0–100) exhibit higher absolute errors, reflecting the model's adjustment phase. Errors significantly reduce between 100–300 cycles, demonstrating improved accuracy as the RNN stabilizes and learns sequential dependencies. Beyond 300 cycles, error trends vary: some instances maintain low error, while others show gradual increases or sharp spikes, particularly after 500 cycles. Key metrics include an MAE of 0.1528 and RMSE of 0.1987, with R-squared at 0.9642. Enhancements like attention mechanisms or hybrid models could improve the RNN's performance in later cycles.

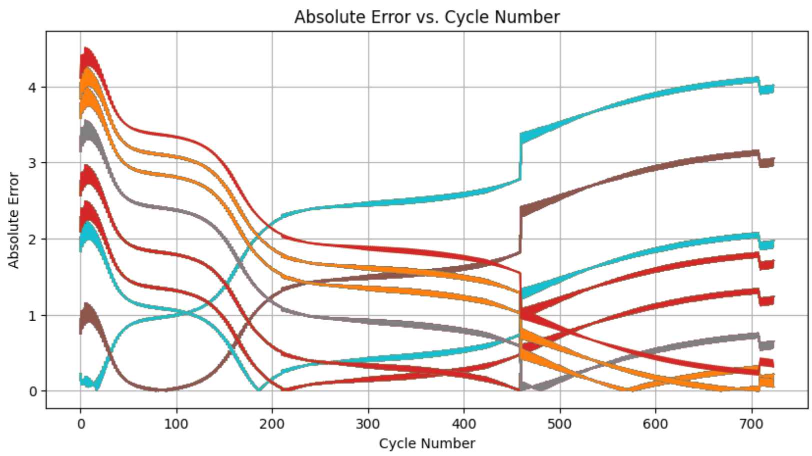

Fig. 20. Absolute Error vs Cycle Number using RNN

RNN demonstrates strong predictive capabilities for SoH, with an MAE of 0.1295, indicating minimal average error between predicted and actual values. The RMSE of 0.1681 reflects consistent performance while managing large errors. An R-squared (R^2) value of 0.9789 shows the model explains 97.89% of the variability in SoH, comparable to LSTM performance. The validation loss of 0.0283 emphasizes the model's ability to generalize with minimal overfitting.

3.1.5 The Best Algorithm amongst the above all?

Among the evaluated models—DNN, LSTM, RNN, and CNN—the LSTM stands out as the best algorithm for predicting the State of Health (SoH), as evidenced in Table 5.1. The LSTM excels across key performance metrics, achieving the lowest MAE of 0.1293, which is comparable to the RNN and slightly better than the DNN, while significantly outperforming the CNN. The LSTM also achieves the lowest RMSE of 0.1680, demonstrating its robustness in handling both small and large prediction errors. Moreover, the LSTM explains 97.9% of the variability in the actual SoH data, as indicated by the highest R-squared (R^2) value of 0.9790, outperforming both the RNN and DNN, and far surpassing the CNN. Additionally, the LSTM shows the lowest validation loss of 0.0282, indicating effective generalization to unseen data without overfitting. While the RNN performs closely, the LSTM's superiority in sequential data modeling makes it the best choice for SoH prediction. The CNN, however, falls short due to higher errors and lower R^2 values.

IV. DISCUSSION

This study investigated battery performance metrics, including terminal voltage, current, capacity, SoC, and SoH, to assess degradation trends and identify key factors influencing battery health. The findings align with or differ from several prior studies on battery aging, management, and prediction methodologies. The observed gradual capacity and SoH degradation trends align with the findings in [1-2], and [11], which emphasized predictable degradation patterns in LiBs due to electrode wear and electrolyte decomposition. These trends also resonate with [46], which documented capacity fade as a result of cycling stress. The strong linear relationships between SoH and capacity, and SoH and charge volt-age, are consistent with [8], which highlighted the utility of such correlations for predictive maintenance.

The study's findings regarding temperature's limited impact within a controlled range align with [26], which noted that thermal effects become significant only under broader temperature variations. Similarly, the stability of terminal voltage despite SoH decline corroborates the conclusions in [6], which identified voltage stability as a key indicator of consistent battery performance. The study's emphasis on advanced da-ta-driven prediction models for SoH, such as those based on deep learning techniques, is consistent with the works in [3] and [39], which advocated for leveraging ML to enhance battery management. The adoption of LSTM networks for time-series predictions aligns with [14] and [27], which demonstrated the effectiveness of LSTMs in capturing temporal dependencies in battery data.

The minimal impact of temperature observed in this study diverges from the findings in [36] and [38], which reported significant temperature-induced degradation under broader environmental conditions. This discrepancy highlights the need for future studies encompassing more diverse temperature ranges to fully understand thermal effects on battery health. The study's performance metrics for CNNs were less robust compared to LSTMs, which contrasts with findings in [29] and [13], which documented high predictive accuracy for CNNs in state-of-health estimation tasks. This variance could be attributed to differences in da-ta preprocessing, feature selection, or CNN architecture design. While the study observed anomalies in current trends during charge-discharge cycles, it did not delve deeply into the causes. In contrast, [33] identified free radicals and other chemical interactions as potential contributors to such deviations. Further chemical analysis could provide insights into these discrepancies.

V. CONCLUSION

LiBs are essential for consumer electronics, EVs, and renewable energy systems. This study aims to predict LiB performance, particularly in EVs, using ML models. Key parameters such as terminal voltage, current, charge voltage, charge current, capacity, SoH, SOC, and temperature were analyzed across multiple charge-discharge cycles. The results showed stable performance under controlled conditions. Terminal voltage fluctuated between 3.65 V and 2.0 V, with current alternating between 1 A during charging and -1 A during discharging, with extreme values around -2 A. Battery capacity decreased from 2.25 Ah to 2.175 Ah, and SoH dropped from 100% to 96%, reflecting natural aging patterns. A strong linear correlation between SoH and capacity was found, suggesting that capacity is a critical predictor of battery health and remaining life. SOC fluctuated between 0% and 100%, with rapid increases above 2.5. However, the study's fixed thermal condition of 25°C limited the ability to explore temperature's broader impact, indicating the need for future studies with varying temperatures. A comparative analysis of four ML models DNN, LSTM, RNN, and CNN showed that LSTM outperformed the others. LSTM achieved a MAE of 0.1293 and RMSE of 0.1680, both indicating minimal prediction deviations. It also had the highest R-squared value of 0.9790, explaining 97.9% of the variance in SoH data, and the lowest validation loss of 0.0282, showing strong generalization. The study highlights the importance of SoH-capacity correlation for proactive maintenance, performance optimization, and battery life extension. Future studies should explore broader temperature ranges, develop advanced thermal management solutions, and consider different battery chemistries. These findings have significant implications for improving battery management systems, optimizing performance, and extending lifespan in EVs and renewable energy systems.

REFERENCES

[1] Marinaro, M., Bresser, D., Beyer, E., Faguy, P., Hosoi, K., Li, H., ... & Passerini, S. (2020). Bringing forward the development of battery cells for automotive applications: Perspective of R&D activities in China, Japan, the EU and the USA. Journal of Power Sources, 459, 228073.

[2] Howey, D. A., Roberts, S. A., Viswanathan, V., Mistry, A., Beuse, M., Khoo, E., ... & Sulzer, V. (2020). Free radicals: making a case for battery modeling. The Electrochemical Society Interface, 29(4), 30.

[3] Liu, K., Hu, X., Zhou, H., Tong, L., Widanage, W. D., & Marco, J. (2021). Feature analyses and modeling of lithium-ion battery manufacturing based on random forest classification. IEEE/ASME Transactions on Mechatronics, 26(6), 2944-2955.

[4] Burzyński, D., & Kasprzyk, L. (2021). A novel method for the modeling of the state of health of lithium-ion cells using machine learning for practical applications. Knowledge-Based Systems, 219, 106900.

[5] Liu, P., Wu, Y., She, C., Wang, Z., & Zhang, Z. (2022). Comparative study of incremental capacity curve determination methods for lithium-ion batteries considering the real-world situation. IEEE Transactions on Power Electronics, 37(10), 12563-12576.

[6] Cahyani, D. E., Setyawan, F. F., Hariadi, A. D., Gumilar, L., & Junoh, A. K. (2023, September). Comparison of Regression Methods for Estimation of State-of-Health in Lithium-Ion Batteries. In 2023 International Conference on Electrical and Information Technology (IEIT) (pp. 202-206). IEEE.

[7] Poh, W. Q. T., Xu, Y., & Tan, R. T. P. (2023, July). Data-Driven Estimation of Li-Ion Battery Health using a Truncated Time-based Indicator and LSTM. In 2023 IEEE Power & Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.

[8] Dos Reis, G., Strange, C., Yadav, M., & Li, S. (2021). Lithium-ion battery data and where to find it. Energy and AI, 5, 100081.

[9] Gur TJIEA, France. Global EV outlook 2020 entering the decade of electric drive. 2020.

[10] Meshram P, Mishra A, Abhilash Sahu R. Environmental impact of spent lithium ion batteries and green recycling perspectives by organic acids - a review. Chemosphere 2020;242:125291.

[11] Xiao J, Li J, Xu Z. Challenges to future development of spent lithium ion batteries recovery from environmental and techno-logical perspectives. Environ Sci Technol 2020;54:9–25.

[12] Fujita T, Chen H, Wang K-t, He C-l, Wang Y-b, Dodbiba G, et al. Reduction, reuse and recycle of spent Li-ion batteries for automobiles: a review. Int J Miner Metall Mater 2021;28:179–92.

[13] Pradhan S, Nayak R, Mishra S. A review on the recovery of metal values from spent nickel metal hydride and lithium-ion batteries. Int J Environ Sci Technol 2022;19:4537–54.

[14] Tran, M. K., DaCosta, A., Mevawalla, A., Panchal, S., & Fowler, M. (2021). Comparative study of equivalent circuit models performance in four common lithium-ion batteries: LFP, NMC, LMO, NCA. Batteries, 7(3), 51.

[15] China Industrial Association of Power (CIAP) Sources: China's top 20 power lithium-ion battery companies in terms of in-stalled capacity in 2020: http://www.ciaps.org. cn/news/show-htm-itemid-37907.html.

[16] Baum ZJ, Bird RE, Yu X, Ma J. Lithium-ion battery Recycling─Overview of techniques and trends. ACS Energy Lett 2022:712–9.

[17] Chen S-P, Lv D, Chen J, Zhang Y-H, Shi F-N. Review on defects and modification methods of LiFePO4 cathode material for lithium-ion batteries. Energy & Fuels; 2022.

[18] Wang, M., Liu, K., Dutta, S., Alessi, D. S., Rinklebe, J., Ok, Y. S., & Tsang, D. C. (2022). Recycling of lithium iron phosphate batteries: Status, technologies, challenges, and prospects. Renewable and Sustainable Energy Reviews, 163, 112515.

[19] Zhao, T., Li, W., Traversy, M., Choi, Y., Ghahreman, A., Zhao, Z., ... & Song, Y. (2024). A review on the recycling of spent lithium iron phosphate batteries. Journal of Environmental Management, 351, 119670.
[20] Yang, X. G., Liu, T., & Wang, C. Y. (2021). Thermally modulated lithium iron phosphate batteries for mass-market electric vehicles. Nature Energy, 6(2), 176-185.
[21] Torregrosa, A. J., Broatch, A., Olmeda, P., & Agizza, L. (2023). A generalized equivalent circuit model for lithium-iron phosphate batteries. Energy, 284, 129316.
[22] Lagnoni, M., Scarpelli, C., Lutzemberger, G., & Bertei, A. (2024). Critical comparison of equivalent circuit and physics-based models for lithium-ion batteries: A graphite/lithium-iron-phosphate case study. Journal of Energy Storage, 94, 112326.
[23] Thelen, A., Lui, Y. H., Shen, S., Laflamme, S., Hu, S., Ye, H., & Hu, C. (2022). Integrating physics-based modeling and machine learning for degradation diagnostics of lithium-ion batteries. Energy Storage Materials, 50, 668-695.
[24] How, D. N., Hannan, M. A., Lipu, M. S. H., Sahari, K. S., Ker, P. J., & Muttaqi, K. M. (2020). State-of-charge estimation of li-ion battery in electric vehicles: A deep neural network approach. IEEE Transactions on Industry Applications, 56(5), 5565-5574.
[25] Tian, J., Xiong, R., Shen, W., & Lu, J. (2021). State-of-charge estimation of LiFePO4 batteries in electric vehicles: A deep-learning enabled approach. Applied Energy, 291, 116812.
[26] Kara, A. (2021). A data-driven approach based on deep neural networks for lithium-ion battery prognostics. Neural Computing and Applications, 33(20), 13525-13538.
[27] Wang, S., Ren, P., Takyi-Aninakwa, P., Jin, S., & Fernandez, C. (2022). A critical review of improved deep convolutional neural network for multi-timescale state prediction of lithium-ion batteries. Energies, 15(14), 5053.
[28] Gu, X., See, K. W., Li, P., Shan, K., Wang, Y., Zhao, L., ... & Zhang, N. (2023). A novel state-of-health estimation for the lithium-ion battery using a convolutional neural network and transformer model. Energy, 262, 125501.
[29] Buchanan, S., & Crawford, C. (2024). Probabilistic lithium-ion battery state-of-health prediction using convolutional neural networks and Gaussian process regression. Journal of Energy Storage, 76, 109799.
[30] Chemali, E., Kollmeyer, P. J., Preindl, M., Fahmy, Y., & Emadi, A. (2022). A convolutional neural network approach for estimation of li-ion battery state of health from charge profiles. Energies, 15(3), 1185.
[31] Zraibi, B., Okar, C., Chaoui, H., & Mansouri, M. (2021). Remaining useful life assessment for lithium-ion batteries using CNN-LSTM-DNN hybrid method. IEEE Transactions on Vehicular Technology, 70(5), 4252-4261.
[32] Wang, Z., Zhao, X., Zhen, D., Pombo, J., Yang, W., Gu, F., & Ball, A. (2024). Adaptable capacity estimation of lithium-ion battery based on short-duration random constant-current charging voltages and convolutional neural networks. Energy, 306, 132541.
[33] Schmitt, J., Horstkötter, I., & Bäker, B. (2023). Electrical lithium-ion battery models based on recurrent neural networks: A holistic approach. Journal of Energy Storage, 58, 106461.
[34] Chen, J., Feng, X., Jiang, L., & Zhu, Q. (2021). State of charge estimation of lithium-ion battery using denoising autoencoder and gated recurrent unit recurrent neural network. Energy, 227, 120451.
[35] Feng, X., Chen, J., Zhang, Z., Miao, S., & Zhu, Q. (2021). State-of-charge estimation of lithium-ion battery based on clockwork recurrent neural network. Energy, 236, 121360.
[36] Liu, Y., Sun, J., Shang, Y., Zhang, X., Ren, S., & Wang, D. (2023). A novel remaining useful life prediction method for lithi-um-ion battery based on long short-term memory network optimized by improved sparrow search algorithm. Journal of Energy Storage, 61, 106645.
[37] Pang, Y., Dong, A., Wang, Y., & Niu, Z. (2024). Deep learning from three-dimensional Lithium-ion battery multiphysics model Part II: Convolutional neural network and long short-term memory integration. Energy and AI, 17, 100398.
[38] Gong, Y., Zhang, X., Gao, D., Li, H., Yan, L., Peng, J., & Huang, Z. (2022). State-of-health estimation of lithium-ion batteries based on improved long short-term memory algorithm. Journal of Energy Storage, 53, 105046.
[39] Ang, E. Y., & Paw, Y. C. (2022). Linear model for online state of health estimation of lithium-ion batteries using segmented discharge profiles. IEEE Transactions on Transportation Electrification, 9(2), 2464-2471.
[40] Athappan, V., Piriyadharshini, D., Suganthi, S., Abimanyu, A., Ranganathan, S., Saravanabalaji, M., & Muthuramalingam, E. (2023, June). Health Monitoring of E-Vehicle Battery Using Machine Learning. In 2023 2nd International Conference on Ad-vancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA) (pp. 1-6). IEEE.
[41] Allirani, S., Pooja, K., Soundarya, E., & Nair, S. S. (2023, July). Li-ion Battery Life Estimation using K-Nearest Neighbor Al-gorithm. In 2023 2nd International Conference on Edge Computing and Applications (ICECAA) (pp. 1606-1610). IEEE.
[42] Vilsen, S. B., & Stroe, D. I. (2024). Dataset of lithium-ion battery degradation based on a forklift mission profile for state-of-health estimation and lifetime prediction. Data in Brief, 52, 109861.
[43] Sheikhani, A., & Agic, E. (2024). Lithium-Ion Battery SOH Forecasting With Deep Learning Augmented By Explainable Machine Learning.
[44] Safavi, V., Mohammadi Vaniar, A., Bazmohammadi, N., Vasquez, J. C., & Guerrero, J. M. (2024). Battery Remaining Useful Life Prediction Using Machine Learning Models: A Comparative Study. Information, 15(3), 124.
[45] Oyucu, S., Doğan, F., Aksöz, A., & Biçer, E. (2024). Comparative Analysis of Commonly Used Machine Learning Approaches for Li-Ion Battery Performance Prediction and Management in Electric Vehicles. Applied Sciences, 14(6), 2306.

[46] Zhang, Y. (2023). Data-driven battery aging diagnostics and prognostics (Master's thesis, Chalmers Tekniska Hogskola (Sweden).

[47] Luo, K., Chen, X., Zheng, H., & Shi, Z. (2022). A review of deep learning approach to predicting the state of health and state of charge of lithium-ion batteries. Journal of Energy Chemistry, 74, 159-173.

[48] Guo, Y., Huang, K., & Hu, X. (2021). A state-of-health estimation method of lithium-ion batteries based on multi-feature extracted from constant current charging curve. Journal of Energy Storage, 36, 102372.

[49] Mocera, F., Somà, A., & Clerici, D. (2020, September). Study of aging mechanisms in lithium-ion batteries for working vehicle applications. In 2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies (EVER) (pp. 1-8). IEEE.

[50] Seok, J., Lee, W., Lee, H., Park, S., Chung, C., Hwang, S., & Yoon, W. S. (2024). Aging Mechanisms of Lithium-ion Batteries. J. Electrochem. Sci. Technol, 15, 51-66.

[51] Yang, S., Zhang, C., Jiang, J., Zhang, W., Zhang, L., & Wang, Y. (2021). Review on state-of-health of lithium-ion batteries: Characterizations, estimations and applications. Journal of Cleaner Production, 314, 128015.

[52] Zhu, J., Xu, W., Knapp, M., Darma, M. S. D., Mereacre, L., Su, P., ... & Ehrenberg, H. (2023). A method to prolong lithium-ion battery life during the full life cycle. Cell Reports Physical Science, 4(7).

[53] Sharma, S. K., Sharma, G., Gaur, A., Arya, A., Mirsafi, F. S., Abolhassani, R., ... & Mishra, Y. K. (2022). Progress in electrode and electrolyte materials: path to all-solid-state Li-ion batteries. Energy Advances, 1(8), 457-510.

[54] Zhang, Y. (2023). Data-driven battery aging diagnostics and prognostics (Master's thesis, Chalmers Tekniska Hogskola (Sweden)).

[55] Thelen, A. C. (2023). Machine learning-based aging models for estimating battery state of health and predicting future deg-radation (Doctoral dissertation, Iowa State University).

[56] Liu, H., Li, C., Hu, X., Li, J., Zhang, K., Xie, Y., ... & Song, Z. (2025). Multi-modal framework for battery state of health evaluation using open-source electric vehicle data. Nature Communications, 16(1), 1137.

[57] Li, K., & Chen, X. (2025). Machine Learning-Based Lithium Battery State of Health Prediction Research. Applied Sciences, 15(2), 516.